

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Hillman, Jennifer L.
Lal, Preeti
Corley, Neil C.
Shah, Purvi

(ii) TITLE OF THE INVENTION: HUMAN CITRATE SYNTHASE HOMOLOG

(iii) NUMBER OF SEQUENCES: 3

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
(B) STREET: 3174 Porter Drive
(C) CITY: Palo Alto
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 94304

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: To Be Assigned
(B) FILING DATE: Herewith
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Billings, Lucy J.
(B) REGISTRATION NUMBER: 36,749
(C) REFERENCE/DOCKET NUMBER: PF-0425 US

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 650-855-0555
(B) TELEFAX: 650-845-4166
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 466 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: ENDCNOT03
(B) CLONE: 2171653

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met Ala Leu Leu Thr Ala Ala Ala Arg Leu Leu Gly Thr Lys Asn Ala
 1 5 10 15
 Ser Cys Leu Val Leu Ala Ala Arg His Ala Ser Ala Ser Ser Thr Asn
 20 25 30
 Leu Lys Asp Ile Leu Ala Asp Leu Ile Pro Lys Glu Gln Ala Arg Ile
 35 40 45
 Lys Thr Phe Arg Gln Gln His Gly Lys Thr Val Val Gly Gln Ile Thr
 50 55 60
 Val Asp Met Met Tyr Gly Gly Met Arg Gly Met Lys Gly Leu Val Tyr
 65 70 75 80
 Glu Thr Ser Val Leu Asp Pro Asp Glu Gly Ile Arg Phe Arg Gly Phe
 85 90 95
 Ser Ile Pro Glu Cys Gln Lys Leu Leu Pro Lys Ala Lys Gly Gly Glu
 100 105 110
 Glu Pro Leu Pro Glu Gly Leu Phe Trp Leu Leu Val Thr Gly His Ile
 115 120 125
 Pro Thr Glu Glu Gln Val Ser Trp Leu Ser Lys Glu Trp Ala Lys Arg
 130 135 140
 Ala Ala Leu Pro Ser His Val Val Thr Met Leu Asp Asn Phe Pro Thr
 145 150 155 160
 Asn Leu His Pro Met Ser Gln Leu Ser Ala Ala Val Thr Ala Leu Asn
 165 170 175
 Ser Glu Ser Asn Phe Ala Arg Ala Tyr Ala Gln Gly Ile Ser Arg Thr
 180 185 190
 Lys Tyr Trp Glu Leu Ile Tyr Glu Asp Ser Met Asp Leu Ile Ala Lys
 195 200 205
 Leu Pro Cys Val Ala Ala Lys Ile Tyr Arg Asn Leu Tyr Arg Glu Gly
 210 215 220
 Ser Gly Ile Gly Ala Ile Asp Ser Asn Leu Asp Trp Ser His Asn Phe
 225 230 235 240
 Thr Asn Met Leu Gly Tyr Thr Asp His Gln Phe Thr Glu Leu Thr Arg
 245 250 255
 Leu Tyr Leu Thr Ile His Ser Asp His Glu Gly Asn Val Ser Ala
 260 265 270
 His Thr Ser His Leu Val Gly Ser Ala Leu Ser Asp Pro Tyr Leu Ser
 275 280 285
 Phe Ala Ala Ala Met Asn Gly Leu Ala Gly Pro Leu His Gly Leu Ala
 290 295 300
 Asn Gln Glu Val Leu Val Trp Leu Thr Gln Leu Gln Lys Glu Val Gly
 305 310 315 320
 Lys Asp Val Ser Asp Glu Lys Leu Arg Asp Tyr Ile Trp Asn Thr Leu
 325 330 335
 Asn Ser Gly Arg Val Val Pro Gly Tyr Gly His Ala Val Leu Arg Lys
 340 345 350
 Thr Asp Pro Arg Tyr Thr Cys Gln Arg Glu Phe Ala Leu Lys His Leu
 355 360 365
 Pro Asn Asp Pro Met Phe Lys Leu Val Ala Gln Leu Tyr Lys Ile Val
 370 375 380
 Pro Asn Val Leu Leu Glu Gln Gly Lys Ala Lys Asn Pro Trp Pro Asn
 385 390 395 400
 Val Asp Ala His Ser Gly Val Leu Leu Gln Tyr Tyr Gly Met Thr Glu
 405 410 415
 Met Asn Tyr Tyr Thr Val Leu Phe Gly Val Ser Arg Ala Leu Gly Val
 420 425 430
 Leu Ala Gln Leu Ile Trp Ser Arg Ala Leu Gly Phe Pro Leu Glu Arg
 435 440 445
 Pro Lys Ser Met Ser Thr Glu Gly Leu Met Lys Phe Val Asp Ser Lys
 450 455 460

Ser Gly
465

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2131 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: ENDCNOT03
- (B) CLONE: 2171653

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CGCCGGTTCG	TCTACTCTTT	CCTTCAGCCG	CCTCCTTCA	ACCTTGTCAA	CCCCTCGGCG	60
CGGCCTCTGG	TGCAGCGCG	GCGGCTCCTG	TTCCTGCCG	AGCTCTCTCC	CTTCTTACC	120
TCCCCACCA	ATCCGGAGA	TCGCCCCCA	TGGCTTACT	TACTGCCGC	GCCCCGGCTCT	180
TGGGAACCA	GAATGCATCT	TGTCTTGTC	TTGCAGCCCC	GCATGCCAGT	GCTTCCTCCA	240
CGAAATTGAA	AGACATATTG	GCTGACCTGA	TACCTAAGGA	GCAGGCCAGA	ATTAAGACTT	300
TCAGGCAGCA	ACATGGCAAG	ACGGTGGTGG	GCCAAATCAC	TGTGGACATG	ATGTATGGTG	360
GCATGAGAGG	CATGAAGGGA	TTGGTCTATG	AAACATCAGT	TCTTGATCCT	GATGAGGGCA	420
TCCGTTTCCG	AGGCCTTAGT	ATCCCTGAAT	GCCAGAAACT	GCTACCCAAG	GCTAAGGGTG	480
GGGAAGAAC	CCTGCCTGAG	GGCTTATT	GGCTGCTGGT	AACCTGGACAT	ATCCCAACAG	540
AGGAACAGGT	ATCTTGGCTC	TCAAAGAGT	GGGCAAAGAG	GGCAGCTCTG	CCTTCCCCTG	600
TGGTCACCAC	GCTGGACAAC	TTTCCCACCA	ATCTACACCC	CATGTCCTCAG	CTCAGTGCAG	660
CTGTTACAGC	CCTCAACAGT	GAAAGTAAC	TTGCCGAGC	ATATGCACAG	GGTATCAGCC	720
GAACCAAGTA	CTGGGAGTTG	ATTATGAAG	ACTCTATGGA	TCTAATCGCA	AAGCTACCTT	780
GTGTTGCAGC	AAAGATCTAC	CGAAATCTCT	ACAGAGAAGG	CAGCGGTATT	GGGGCCATTG	840
ACTCTAACCT	GGACTGGTCT	CACAATTTC	CCAACATGTT	AGGCTATACT	GATCATCAGT	900
TCACTGAGCT	CACGCCCTG	TACCTCACCA	TCCACAGTGA	CCATGAGGGT	GGCAATGTAA	960
GTGCCCATGAC	CAGCCATTG	GTGGGAGTG	CCCTTTCCGA	CCCTTACCTG	TCCCTTGCAG	1020
CAGCCATGAA	CGGCCCTGGCA	GGGCCCTCTCC	ATGGACTGGC	AAATCAGGAA	GTGCTTGTCT	1080
GGCTAACACA	GCTGAGAG	GAAGTTGGCA	AAGATGTGTC	AGATGAGAA	TTACGAGACT	1140
ACATCTGGAA	CACACTAAC	TCAGGACGGG	TTGTTCCAGG	CTATGGCCAT	GCAGTACTAA	1200
QGAAGACTGA	TCCCGATAT	ACCTGTCAGC	GAGAGTTTGC	TCTGAAACAC	CTGCCATTATG	1260
ACCCCATGTT	TAAGTTGGTT	GCTCAGCTGT	ACAAGATTGT	GCCCAATGTC	CTCTTAGAGC	1320
AGGGTAAAGC	CAAGAATCCT	TGGCCAATG	TAGATGCTCA	CAGTGGGGTG	CTGCTCCAGT	1380
ATTATGGCAT	GACGGAGATG	AATTACTACA	CGGTCTGTT	TGGGGTGTCA	CGAGCATTGG	1440
GTGTACTGGC	ACAGCTCATC	TGGAGCCGAG	CCTTAGGCTT	CCCTCTAGAA	AGGCCCAAGT	1500
CCATGAGCAC	AGAGGGTCTG	ATGAAGTTG	TGGACTCTAA	GTCAGGGTAA	AACTGGAGAC	1560
TGGGTGAAAG	TGACTACCAG	AAAGTGAGGA	AGCCTAAATA	AAAAGTATAAC	TTTTGTTCA	1620
GGGGGCCTTT	AAAGACTTAA	GATTAAATT	TATCTGAGGC	ACTGATAATA	TGTTTGAGGT	1680
TAAAATATAA	ATTAAGACTT	TAAAAGATGA	AAAATGGTCC	CTTCTTCCCT	AATCAGCTCC	1740
CTTCCCTCTGC	CTGGTATGAG	TTGCCCATCA	TACGCATGGT	CCTGGAGGAT	GACCAGGACT	1800
AATGCATGTG	CTATGAGTAG	TTTGGCCCC	CTCACTATCT	CTAGAGTGTAG	AATCTGGCTC	1860
CTGTTTCCAT	GGGTCAAAGC	CGGTGCGAGA	GAATCTGTAG	TCACCTTGG	GCTTTAGCTT	1920
CTCTGCCAAG	CCCTCAATAA	GCCAGCAAAC	CAGGACTCTG	CCCTTCTCTGT	TTCCCATAGGA	1980
ATCATGTTGG	ATAGTCAGCT	GTACCAAGCC	CCTTGGCCCT	CICCCATGCA	CACAAACACC	2040
TCCTAGCAAG	ACCTGTTGGT	TAGCTGGACA	TGCTTGGCA	ATTTTTTTAT	ACTACCAAGT	2100
GACCATATTG	GCATGGCATT	TTTGGTGAT	G			2131

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 464 amino acids
- (B) TYPE: amino acid

(C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
 (A) LIBRARY: GenBank
 (B) CLONE: 164419

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Ala Leu Leu Thr Ala Ala Ala Arg Leu Phe Gly Ala Lys Asn Ala
 1 5 10 15
 Ser Cys Leu Val Leu Ala Ala Arg His Ala Ser Ala Ser Thr Asn
 20 25 30
 Leu Lys Asp Ile Leu Ala Asp Leu Ile Pro Lys Glu Gln Ala Arg Ile
 35 40 45
 Lys Thr Phe Arg Gln Gln His Gly Asn Thr Val Val Gly Gln Ile Thr
 50 55 60
 Val Asp Met Met Tyr Gly Gly Met Arg Gly Met Lys Gly Leu Val Tyr
 65 70 75 80
 Glu Thr Ser Val Leu Asp Pro Asp Glu Gly Ile Arg Phe Arg Gly Tyr
 85 90 95
 Ser Ile Pro Glu Cys Gln Lys Met Leu Pro Lys Ala Lys Gly Gly Glu
 100 105 110
 Glu Pro Leu Pro Glu Gly Leu Phe Trp Leu Leu Val Thr Gly Gln Ile
 115 120 125
 Pro Thr Glu Glu Gln Val Ser Trp Leu Ser Lys Glu Trp Ala Lys Arg
 130 135 140
 Ala Ala Leu Pro Ser His Val Val Thr Met Leu Asp Asn Phe Pro Thr
 145 150 155 160
 Asn Leu His Pro Met Ser Gln Leu Ser Ala Ala Ile Thr Ala Leu Asn
 165 170 175
 Ser Glu Ser Asn Phe Ala Arg Ala Tyr Ala Glu Gly Ile His Arg Thr
 180 185 190
 Lys Tyr Trp Glu Leu Ile Tyr Glu Asp Cys Met Asp Leu Ile Ala Lys
 195 200 205
 Leu Pro Cys Val Ala Ala Lys Ile Tyr Arg Asn Leu Tyr Arg Glu Gly
 210 215 220
 Ser Ser Ile Gly Ala Ile Asp Ser Lys Leu Asp Trp Ser His Asn Phe
 225 230 235 240
 Thr Asn Met Leu Gly Tyr Thr Asp Ala Gln Phe Thr Glu Leu Met Arg
 245 250 255
 Leu Tyr Leu Thr Ile His Ser Asp His Glu Gly Asn Val Ser Ala
 260 265 270
 His Thr Ser His Leu Val Gly Ser Ala Leu Ser Asp Pro Tyr Leu Ser
 275 280 285
 Phe Ala Ala Ala Met Asn Gly Leu Ala Gly Pro Leu His Gly Leu Ala
 290 295 300
 Asn Gln Glu Val Leu Val Trp Leu Thr Gln Leu Gln Lys Glu Val Gly
 305 310 315 320
 Lys Asp Val Ser Asp Glu Lys Leu Arg Asp Tyr Ile Trp Asn Thr Leu
 325 330 335
 Asn Ser Gly Arg Val Val Pro Gly Tyr Gly His Ala Val Leu Arg Lys
 340 345 350
 Thr Asp Pro Arg Tyr Thr Cys Gln Arg Glu Phe Ala Leu Lys His Leu
 355 360 365
 Pro His Asp Pro Met Phe Lys Leu Val Ala Gln Leu Tyr Lys Ile Val
 370 375 380
 Pro Asn Val Leu Leu Glu Gln Gly Lys Ala Lys Asn Pro Trp Pro Asn
 385 390 395 400
 Val Asp Ala His Ser Gly Val Leu Leu Gln Tyr Tyr Gly Met Thr Glu

PF-0425-1 DIV

405	410	415
Met Asn Tyr Tyr Thr Val Leu Phe Gly Val Ser Arg Ala Leu Gly Val		
420	425	430
Leu Ala Gln Leu Ile Trp Ser Arg Ala Leu Gly Phe Pro Leu Glu Arg		
435	440	445
Pro Lys Ser Met Ser Thr Asp Gly Leu Ile Lys Leu Val Asp Ser Lys		
450	455	460